SCHS STUDIES

A SPECIAL REPORT SERIES BY THE N.C. DEPARTMENT OF HUMAN RESOURCES, DIVISION OF HEALTH SERVICES, STATE CENTER FOR HEALTH STATISTICS, P.O. BOX 2091, RALEIGH, N.C. 27602

No. 20 July 1981

THE USE OF HOSPITAL DISCHARGE DATA FOR ESTIMATING MORBIDITY IN NORTH CAROLINA

In October of 1978, data were collected on North Carolina residents discharged from short-stay general hospitals. This project involved the cooperation of health planning agencies and the hospital association in North Carolina. As stated in a letter from the N.C. Hospital Association requesting the participation of hospitals, the purpose of the study was " to provide hospitals and health planning agencies some basic data needed for cooperative planning, and to test the feasibility and usefulness of a continuing hospital data base." The main purpose of this report is to present morbidity estimates derived from the October 1978 data, but first it will be useful to provide some background concerning how the study was carried out.

Study Background

There has been discussion for a number of years of the desirability of a hospital "data consortium" in North Carolina. Under this concept, hospitals throughout the state would submit data a single time to a data processor or "broker," which would then act as a clearinghouse for the various and often duplicative information requests received by the hospitals. While good in theory, this idea has proved difficult to implement because of political, financial, and other obstacles. The October 1978 study was designed to demonstrate the feasibility of using a single data submission for multiple purposes. Participating hospitals that completed October abstracting by a certain date were provided a summary report of information required for the geographic service area section of the annual hospital licensure form. In addition, all participating hospitals received patient origin reports oriented toward facility planning, including comparable utilization information for other institutions serving the same geographic area. The central idea was to promote cooperative planning among hospitals and with the health planning agencies, which received the same information. Over half of the participating hospitals took advantage of the licensure report service in spite of the short data turnaround time required. Though this one-month pilot study has not directly led to a centralized hospital data base, it has demonstrated the feasibility of collecting discharge data on a large scale, and has sensitized hospitals and other parties to the advantages of pooling and sharing data.

All but 12 of the 132 short-term general hospitals in the state participated in the October 1978 study. Emphasis was on collecting data through existing systems, and information for about 75 percent of N.C. hospital patients was obtained from discharge abstracts of the ongoing Professional Activity Study (PAS). Another 10 percent of the month's discharges were captured through the computer systems of five large hospitals in the state, with data on the other patients in the study manually coded and then keypunched. Most of the major out-of-state hospitals receiving N.C. patients agreed to code data for N.C. residents discharged in October, and overall about 94 percent of N.C. residents going to short-term general hospitals during the month were captured in the study (1). About 71,000 N.C. residents plus out-of-state residents discharged from N.C. hospitals yielded a total sample size of 73,307. For each patient the following

items of information were collected: hospital ID; admission and discharge dates; age; sex; race; hospital service category; primary, secondary, and tertiary diagnoses; payment source; discharge status; and zip code of patient's residence. Only about half of these variables were employed in the initial planning reports provided to the hospitals, and there remains from the study a substantial body of data with potential for further analysis.

Representativeness of the Sample

A question naturally arises concerning the representativeness of one month's information. To address this question, data from the October study were compared to fiscal year 1978-1979 information derived from the N.C. Division of Facility Services (DFS) annual licensure form. According to the DFS data, 855,132 patients were discharged from 132 North Carolina short-stay general hospitals during the year, including out-of-state patients but excluding newborns and prematures. Multiplying the comparable October study figure times 12 and adjusting for six percent non-participation yields an annual estimate of 840,369 discharges, only 1.7 percent less than the actual figure.

As might be expected, estimates are less accurate when various subdivisions of discharges are considered. In general, a one-month sample is least representative of elective procedures and better represents more random admissions. For example, for the 132 general hospitals, the estimate from the October study of an annual number of thoracic surgery procedures was 9,052, while the number reported through DFS was 6,329, for a difference of 43 percent (using 6,329 in the denominator). For obstetric discharges, on the other hand, the October estimate of 94,640 discharges per year compared to the DFS figure of 89,969, a difference of only 5.2 percent. For other categories where comparison between the two data sets was feasible, the annual estimate from the October data, the DFS figure, and the percent difference between the two are shown below.

	October Estimate	DFS	Percent Difference*
Psychiatry	25,792	18,572	39
General Surgery	118,918	136,945	-13
Neurosurgery	12,922	15,874	-19
Gynecology	62,979	60,960	3
Urology	51,419	48,338	6
Orthopedics	71,279	62,614	14

*DFS figure as the base

In the case of psychiatry, some of the difference may be due to a reporting inconsistency, where only discharges from designated psychiatric units of the general hospitals are reported to DFS while any discharge with a psychiatric diagnosis is counted in the October data.

Clearly, for specific categories of hospital discharges, a one-month sample will not always give accurate estimates of absolute yearly levels of the phenomenon of interest. Such a sample is more appropriate for determining relative differences among population subgroups with regard to pertinent variables.

Some General Results from the October Study

For the 73,307 patients in the October 1978 study (which includes about 9 percent newborns), 60 percent were female and 40 percent were male. Seventy-seven percent were reported as white, 20 percent black, 1 percent Asiatic and other nonwhite, and 2 percent had no race reported. Eighty-eight percent of these patients were discharged to their residence, 1.5 percent were transferred to another hospital, 1.4 percent were transferred to a nursing home, 0.1 percent were transferred to home health care, 2.5 percent were discharged dead, and for 6.5 percent, discharge status was "other" or not recorded. Average age of patients varies considerably by type of hospital service. Average length of hospital stay was 7.0 days, with a range from 1.5 days for abortions to 13.4 days for thoracic surgery. Data for expected source of payment are shown in Table 1.

Table 1

Percent of Patients by Source of Payment: North Carolina, October 1978 Short-Stay General Hospitals

Medicare	21.8
Workmen's Compensation	1.2
Medicaid	6.2
Other Government	3.0
Blue Cross	22.5
Commercial Insurance	30.1
Self Pay	10.3
Charity	0.7
Not Listed Above	4.1

With data collected at the level of the individual patient, it is of course possible to look at many different breakdowns and crosstabulations of the information. We could examine, for example, payment source by age, discharge status by race, or hospital service by sex by length of stay. Any of these analyses could be carried out for regions, counties, or individual hospitals or zip codes. Requests that have been answered using this data set, in addition to the original tabulations that were produced for all hospitals, include a distribution of patients by race and sex by hospital within a Health Service Area (HSA) and a distribution of cancer patients by hospital by county of origin for residents of another HSA. Other useful tabulations for planning and other purposes could also be derived.

Estimating Morbidity

While we currently have rather extensive information concerning fertility and mortality of North Carolinians, comparatively little is known about morbidity or sickness in the population. Communicable disease data are reported through the Division of Health Services, but reportable communicable diseases are a small subset of total morbidity and there are thought to be substantial problems with completeness of

reporting for some of these diseases. The North Carolina Citizen Survey asks a few health-related questions, but with a limited sample size, estimates below the state level may not be stable. Besides a full-scale, very expensive health interview and examination survey, the major remaining way to obtain systematic morbidity information is through hospital discharge data. Though only sickness that is treated in the hospital is reflected in these data, this component is very important for assessing the health status of a population.

In 1978, general hospitals in North Carolina were operating with two disease coding systems: ICDA-8 and H-ICDA-2. The H system, employed by all PAS hospitals and some others, is an adaptation of ICDA-8 for hospital usage developed by the Commission on Professional and Hospital Activities (CPHA) in Ann Arbor, Michigan. For large categories such as heart disease or cancer, the two systems yield almost identical results, but for more specific diagnoses there may be significant differences in coding. To facilitate the present and also future analysis of the discharge data, diagnoses for all hospitals using H-ICDA-2 were converted to the corresponding ICDA-8 code using a conversion tape purchased from CPHA. As a result, diagnosis data from the two types of hospitals can be pooled for analysis, and the data become more comparable to mortality coding, which employed ICDA-8 for 1978 events.

Discharge Rates for Leading Causes of Mortality

The State Center for Health Statistics (SCHS) regularly publishes mortality rates for the leading underlying causes of death, so it seemed natural to compute discharge rates for ten of these diagnostic categories. This will allow, as a point of departure, comparisons for categories familiar to most readers of SCHS publications. Four categories--suicide, homicide, motor vehicle accidents, and other accidents--were omitted because of a difference between morbidity and mortality coding. Accidents, poisonings, and violence are classified both under nature of injury (e.g., contusion of the brain) and under external cause (e.g., motor vehicle accident). External causes are always chosen as the underlying cause of death, with the nature of injury treated as a mentioned condition on the certificate. On the hospital medical record, however, the nature of injury code is always chosen as the final diagnosis (first listed), and the external cause is coded as an additional diagnosis on the record. So it is not possible to compare hospital and death certificate coding for externally-caused conditions by looking only at the first listed diagnosis.

Table 2 displays, for the state, general hospital discharge rates based on final diagnosis explaining admission and death rates based on underlying cause for ten diagnostic categories. The discharge rates refer to N.C. residents, not facilities, and the October data have been multiplied by twelve and adjusted for missing hospitals (6 percent of the patients missing for the state as a whole) to yield the annualized rates shown here. Among the ten diagnoses studied, heart disease is the most frequent cause of hospital admission, followed by cancer, cerebrovascular disease, influenza and pneumonia, chronic obstructive lung disease, and diabetes. The ten causes in Table 2 account for about 18 percent of all hospital discharges. The total rate shown here includes discharges of newborns from the hospital.

Heart disease, cancer, and cerebrovascular disease are the top three causes of death. Comparing discharge and death rates reveals that hospital episodes for each of the disease categories occur much more frequently than deaths. Overall, there are about 19 general hospital discharges for every North Carolina resident death, and this ratio for the state as a whole varies from 32 for hypertension as the primary diagnosis to 3 for heart disease as the primary diagnosis. Other causes with a high discharge-to-death ratio are diabetes, influenza and pneumonia, chronic obstructive lung disease,

and nephritis and nephrosis, reflecting a relatively higher level of disease in the population than death data alone would indicate. For most of these causes a high discharge-to-death ratio is understandable since they are long-term contributors to the chronic disease process resulting in death. The ten causes shown in Table 2 account for 77 percent of all deaths in North Carolina.

Table 2

General Hospital Discharges (Annualized) and Deaths
Per 100,000 Resident Population for North Carolina, 1978:

Based on Final Diagnosis Explaining
Admission and Underlying Cause

	Rate	
	Discharge	Death
Total	16,266	848
Cancer	676	160
Diabetes	188	14
Heart Disease	972	313
Hypertension	95	3
Cerebrovascular Disease	295	84
Arteriosclerosis	82	11
Influenza/Pneumonia	284	26
Chronic Obstructive Lung Disease	250	22
Cirrhosis of the Liver	38	12
Nephritis/Nephrosis	49	5
All Other Diagnoses	13,337	198

Multiple Conditions Analysis

In the October discharge study three diagnoses were collected for each patient record: final diagnosis explaining admission and first and second additional diagnoses (if any). The medical record may contain more than three diagnoses, but only the first three were captured here. In comparing discharges and deaths based on a single diagnosis, final diagnosis explaining admission should be compared to underlying cause of death, though the two are in fact only roughly parallel. An analysis of October 1978 general hospital deaths comparing discharge data with the Division of Health Services' death file indicated some inconsistencies in coding. For example, while cerebrovascular disease was coded with roughly equal frequency as final diagnosis or underlying cause for deaths in the two data files, heart disease appeared nearly two times more often as an underlying cause in the death file than it did as final diagnosis explaining admission among the October patients discharged dead (adjusting for missing hospitals). Presumably many deaths in the hospital with heart disease coded as a secondary or tertiary diagnosis get coded in the death file with heart disease as the underlying cause of death. To promote a better comparison of the discharge and death data, the rates in Table 2 were recomputed using all of the diagnoses coded on the hospital and death records.

In looking at disease prevalence in terms of multiple conditions, any mention of a condition in any of the diagnostic categories is counted as a case. On the death

record there may be up to fourteen conditions mentioned, one of which is usually chosen as the underlying cause, while with the October discharge data three diagnoses are examined and counted. Thus, records may be counted more than once under this method; for example, a heart disease death (underlying cause) with hypertension and diabetes mentioned on the certificate would be counted as a case for each of the three diagnoses. Table 3 displays general hospital discharge rates and Table 4 displays death rates, based on multiple conditions present, for residents of the six HSAs and for the state. The discharge data are again annualized. Inferences below the HSA level based on the hospital discharge data would be precarious in some areas because of missing facilities.

Heart disease and cancer still show up as the major diagnoses in Tables 3 and 4, but other diagnoses are much more prevalent when multiple conditions are considered than they are in Table 2. For the state as a whole, among these ten causes, diabetes shows up in Table 3 as the third most common discharge diagnosis, at a level more than three times as great as the rate based on final diagnosis alone (Table 2). Hypertension is fourth, and the rate is more than five times that in Table 2. Arteriosclerosis shows up much more frequently as a mentioned condition than as a final diagnosis explaining admission. Comparing Tables 4 and 2 for the state as a whole, diabetes, hypertension, and arteriosclerosis again show up much more often when mentioned conditions are considered in addition to underlying cause. HSA I in many cases has the highest discharge and death rates for these causes, though this HSA has the highest percent over age 65 in the state and age-adjusted rates would show less of a difference. Nevertheless, this HSA has a significantly higher degree of morbidity to which the medical care system must adapt. Further analysis of these tables by the reader should reveal other patterns of interest.

Comparison of Morbidity Measures

In comparing discharge and death data based on multiple conditions, a question arises concerning the extent to which these two types of data are measuring different aspects of morbidity. In an attempt to answer this question, correlation coefficients were computed between discharge and death rates across the six HSAs for each of the ten diagnoses. With only six data points, a correlation coefficient of less than .82 is not statistically significant at the 95 percent confidence level. For only two of the diagnoses, cancer and arteriosclerosis, were the coefficients greater than .82, indicating that in general we cannot say with 95 percent certainty that discharge and death rates measure the same thing. In fact, for cirrhosis of the liver, the two sets of rates correlated at $r = -.69$. Discharge and death rates were then each correlated with disease prevalence estimates by HSA from the N.C. Citizen Survey for seven of the diagnoses. These correlations were also generally not significant, with only the coefficients between the citizen survey estimates and discharge rates for cerebrovascular disease and chronic obstructive lung disease being .82 or greater. For most diagnoses, however, the discharge rates correlated higher with the Citizen Survey estimates than did the death rates. These findings suggest that hospital discharge data may be a better measure of general morbidity than multiple conditions mortality data.

Discharge Rates for Other Diagnoses

Though it is of some interest to compare hospital discharge rates to death rates, the leading death diagnoses account for less than one-fifth of hospital discharges. To examine what disease categories do account for hospital morbidity, discharge rates were

Table 3

General Hospital Discharges Per 100,000
Resident Population by HSA, 1978 (Annualized):

Based on Diseases Mentioned in Any of Three Diagnoses

	HSA I	HSA II	HSA III	HSA IV	HSA V	HSA VI	N.C.
Cancer	867	817	783	744	700	777	785
Diabetes	621	626	513	537	681	658	608
Heart Disease	2,019	1,698	1,424	1,401	1,706	1,647	1,658
Hypertension	511	513	356	480	573	584	504
Cerebrovascular Disease	620	403	360	389	462	454	449
Arteriosclerosis	534	282	253	301	318	299	331
Influenza/Pneumonia	577	316	382	308	347	336	379
Chronic Obstructive Lung Disease	552	384	343	398	371	466	421
Cirrhosis of the Liver	78	88	54	67	80	76	74
Nephritis/Nephrosis	74	71	73	67	86	96	78

Table 4

Deaths Per 100,000 Resident Population by HSA, 1978:

Based on Multiple Conditions Present at Death

	HSA I	HSA II	HSA III	HSA IV	HSA V	HSA VI	N.C.
Cancer	195	185	179	172	166	186	181
Diabetes	58	60	58	57	55	70	60
Heart Disease	473	442	434	441	404	435	439
Hypertension	65	73	76	80	68	91	76
Cerebrovascular Disease	152	140	134	132	139	151	142
Arteriosclerosis	182	144	129	138	123	148	145
Influenza/Pneumonia	101	80	70	79	71	77	80
Chronic Obstructive Lung Disease	63	54	56	52	48	56	55
Cirrhosis of the Liver	14	16	20	18	18	18	18
Nephritis/Nephrosis	11	11	13	15	15	17	14

Table 5

General Hospital Discharges Per 100,000
Resident Population for North Carolina 1978 (Annualized):
Based on Final Diagnosis Explaining Admission

	Chapter and Code	Rate
I.	Infective and Parasitic Diseases (000-136)	533
II.	Neoplasms (140-239)	978
III.	Endocrine, Nutritional, and Metabolic Diseases (240-279)	310
IV.	Diseases of the Blood Forming Organs (280-289)	109
V.	Mental Disorders (290-315)	540
VI.	Diseases of the Nervous System and Sense Organs (320-389)	645
VII.	Diseases of the Circulatory System (390-458)	1,787
VIII.	Diseases of the Respiratory System (460-519)	1,217
IX.	Diseases of the Digestive System (520-577)	1,612
X.	Diseases of the Genitourinary System (580-629)	1,608
XI.	Complications of Pregnancy, Childbirth, and the Puerperium (630-678)	1,841
XII.	Diseases of the Skin and Subcutaneous Tissue (680-709)	207
XIII.	Diseases of the Musculoskeletal System and Connective Tissue (710-738)	722
XIV.	Congenital Anomalies (740-759)	132
XV.	Certain Causes of Perinatal Morbidity and Mortality (760-779)	28
XVI.	Symptoms and Ill-Defined Conditions (780-796)	874
XVII.	Accidents, Poisonings, and Violence (800-999)	1,396

NOTE: Total N.C. population for 1978 was used in the denominator for all of the rates calculated in this table in order to promote comparison among these rates and with hospital discharge rates published by the National Center for Health Statistics, though less than the total population is "at risk" for categories XI, and XV.

computed for each of the seventeen major divisions or "chapters" of the ICDA system. These rates, based on final diagnosis explaining admission, are shown in Table 5 for the state as a whole. HSA rates are available upon request. The seventeen categories account for 89 percent of all hospital discharges, the remainder being primarily newborns and other codes representing special conditions and examinations without sickness. The highest discharge rate is for the category of complications of pregnancy, etc. (which includes normal deliveries and therefore is somewhat of a misnomer), followed by diseases of the circulatory system, digestive system, genitourinary system, then accidents, poisonings, and violence, and diseases of the respiratory system. Keep in mind that the rate for mental disorders is understated, since freestanding short-stay psychiatric hospitals and state mental hospitals were not included in the October study. The data could, of course, be broken out by other disease categories as well, but these seventeen standard groupings were considered a logical place to start.

Racial Comparisons

Tabulations produced for a request from the Health Promotion Branch of the Division of Health Services illustrate further uses of hospital discharge data. The task was to compare whites and nonwhites with regard to heart disease, cerebrovascular disease, hypertension, and diabetes. Table 6 shows that while nonwhites had a lower rate of discharge from the hospital for heart and cerebrovascular diseases, their hospital stay was longer and their chance of dying in the hospital was greater (2).

Table 6

Hospitalization Data for N.C. Residents by Race, October 1978 (Rates Annualized)

Primary Diagnosis	Number Hospital Discharges Per 100,000 Population		Average Length of Stay (Days)		Percent Discharged Dead	
	White	Nonwhite	White	Nonwhite	White	Nonwhite
Heart Disease	1,056	651	9.81	10.43	8.5	10.0
Cerebrovascular Disease	300	271	12.28	13.94	15.8	18.0
Hypertension	75	154	6.45	6.95	2.1	3.3
Diabetes	162	258	8.30	10.03	2.3	0.8*
All Hospital Discharges	16,196	14,800	6.99	7.05	2.5	2.5

*based on only 2 deaths

For hypertension and diabetes as the primary diagnosis, nonwhites were discharged from the hospital at almost twice the rate of whites, and here also the average length of stay and percent discharged dead were generally greater for nonwhites. Other tabulations of the data show that 19.5 percent of nonwhites with cerebrovascular disease as the primary diagnosis had hypertension listed as a secondary or tertiary diagnosis, compared to 10.4 percent for whites. Nonwhites also had a higher incidence of hypertension mentioned in conjunction with heart disease, (8.1 percent versus 5.8 percent for whites) and a higher incidence of diabetes mentioned in conjunction with heart disease and with cerebrovascular disease.

As discussed before, another way to look at disease prevalence is to count as a case any mention of a condition in any of the three diagnostic categories. Table 7 shows the results by race for four diagnoses and one combination. Prevalence is much higher than that reflected in Table 6, and about five times as high for hypertension. The racial differentials still hold. Since nonwhites generally have less access to the health care system than whites, the patterns shown here probably understate a presumed greater incidence of hypertension and diabetes among nonwhites. Also, among those people who do reach the hospital, disease incidence shown here reflects only those conditions that get coded onto the medical record, and so there is some degree of understatement. For hypertension in particular there would seem to be understatement, since the primary diagnosis might be "hypertensive heart disease," which would be recorded as heart disease, but a separate hypertension code may not be put on the record. In sum, these data better indicate relative differences than absolute levels of disease.

Table 7

Hospital Discharges Per 100,000 Population:

Disease Mentioned in Any of Three Diagnoses
North Carolina, 1978 (Annualized)

	White	Nonwhite
Heart Disease	1,768	1,252
Cerebrovascular Disease	462	408
Hypertension	449	685
Diabetes	556	760
Hypertension and Diabetes	56	133

Discussion

While the preceding has shown several ways that hospital discharge data can be used to estimate morbidity, it should be reiterated that the results cannot be generalized to non-hospitalized patients. Patients choose or are referred to hospitals in a nonrandom fashion, and disease patterns among the hospitalized do not simply reflect morbidity in the community at large (3,4). As well as having more serious diseases, hospitalized patients tend to have more disease combinations than non-hospitalized patients. Hospital data should not be taken to indicate the general incidence of a disease, except for those diseases where virtually all patients with the disease require hospitalization. However, hospital data should be a better measure of disease incidence than death data. The above breakdown of diagnosis data by race suggests that nonwhites have a higher incidence of hypertension and diabetes, but other factors could account for some of the difference. For example, home and other outpatient care may be less likely for nonwhites, who on the average are of lower socioeconomic status, and this could produce more hospitalization without more disease. The major point is that hospital data represent only a part of the total morbidity picture and should be used with care when making inferences about sickness in the population.

A potential problem with the data used in this report is that multiple hospitalizations by the same person are reflected in the annual estimates. The Capital

Area (N.C.) Professional Standards Review Organization (PSRO) has calculated that 22 percent of federal hospital admissions in their area over the course of a year are readmissions by the same person, though this varies from 10 to 45 percent for individual hospitals. Watts and Acheson (5) found in 1964 in Britain that about half of the hospital admissions for cancer of the cervix uteri during a year were readmissions of the same patient for the same condition, although the average for the nineteen diagnoses they examined was 24 percent. Readmissions during one month are few, but multiplying by 12 to get an annual estimate in effect incorporates this double counting of patients. For some purposes, such as measuring the demand for hospital services as a basis for facility planning, all admissions should be counted. Including readmissions in a discharge rate would produce a weighted rate that might be desirable if it were assumed that multiple hospital episodes represented the more serious cases. For other purposes, however, it would be best to compute discharge rates based on cases or persons excluding readmissions during the year for the same condition. A year's data could be unduplicated for patients readmitted to the same facility using a hospital patient number.

Despite the inability to generalize about total morbidity from hospital data and the potential distortion from multiple admissions by the same patient, discharge data capture a central component of the morbidity picture in North Carolina and are relatively easy to obtain. The ongoing National Hospital Discharge Survey regularly produces profiles of hospital morbidity for the U.S. as a whole (6), but except for isolated studies (7,8) there has been no comparable information for North Carolina. The present report is intended to begin filling this gap.

Future Directions

Though other worthwhile analyses can certainly be derived from this data set, data for 1978 are rapidly becoming out of date, and one month's information has its limitations. October was selected because it corresponded to the month for which patient origin information was asked by the state licensure division, because this month and April are reputed to be the most representative months if only one is chosen, and because for a pilot study it was considered prudent to start relatively small. Also, the manual coding of patient records for more than one month would have been very time consuming for the non-automated hospitals that participated. A follow-up study could, however, reasonably collect more than a month's information with substantial completeness.

The difficulty with arriving at a statewide picture from hospital discharge data is that the information must be pieced together from several different sources. The October study approached the problem by compiling data for institutions: PAS hospitals, hospitals with computers, and manual coding for the others. One could also assemble data from payment sources like Blue Cross/Blue Shield, Medicare and Medicaid and approach private insurance companies, though self-pay patients and other small components would most probably have to be left out. Experience from the October study dictates two conclusions as to the best approach for any future study: a) the PAS data base should be used as the basic building block since about 75 percent of the discharges are available here in a standard format from a single source (9), and b) a full year of information should be collected. A year's PAS data can be purchased for approximately $10,000. Such a purchase is contingent on a data release agreement from each PAS hospital. To the PAS data could be added data from the six large hospitals in the state with their own computerized medical record systems. Professional Standards Review Organization information on all federal patients (Medicare, Medicaid, etc.) can be obtained from the federal government on computer tape for the remaining hospitals. Thus all patients except nonfederal patients in small hospitals not on PAS would be represented in the system, well over 90 percent of the total, without any additional

forms to be filled out by any hospital. The main objective should be to rely as completely as possible on existing data systems.

Even if all of these sources could not be assembled, for financial or other reasons, PAS data alone represent reasonably well the experience of all patients in North Carolina. Comparisons of different subsets of the October 1978 data to the total showed that PAS patients are distributed much like all patients with regard to age, race, sex, hospital service, and discharge status. Blue Cross/Blue Shield and other single sources are much less representative of the whole.

A one-year hospital discharge data collection project would provide an update to the October 1978 study and establish a baseline for periodic future assessments of morbidity and hospital usage in North Carolina.

FOOTNOTES AND REFERENCES

(1) For information on N.C. residents going to bordering out-of-state hospitals see UNC Area Health Education Centers Program, Inpatient Utilization of General Hospitals in North Carolina, October 1974. Coverage within North Carolina was determined using hospital data from the N.C. Division of Facility Services.

(2) The use of statistical significance tests in order to generalize from October to the whole year with a known probability of being correct is not entirely appropriate here since October patients are not a true random sample of the year's patients. The differences in Table 6 may be considered "real" differences for the month of October, 1978, though one could treat these patients as one of many possible measurements of October 1978 and make statistical inferences about the real values based on variation from sample to sample due to measurement error.

(3) National Center for Health Statistics, "Use of Hospital Data for Epidemiologic and Medical-Care Research," Vital and Health Statistics, Series 4, No. 11, June 1969.

(4) David L. Sackett, "Bias in Analytic Research," Journal of Chronic Diseases, Volume 32, 1979, p. 51-63.

(5) S. P. Watts and E. D. Acheson, "Computer Method for Deriving Hospital Inpatient Morbidity Statistics Based on the Person as the Unit," British Medical Journal, 4, 1967, p. 476-477.

(6) See, for example, National Center for Health Statistics, "Utilization of Short-Stay Hospitals: Annual Summary of the United States, 1978," Vital and Health Statistics, Series 13, No. 46, 1980.

(7) Division of Health Affairs, University of N.C., Patient Characteristics and Treatment in N.C. Hospitals in 1968: An Analysis of PAS Data for N.C. Hospitals, August 1971.

(8) N.C. Division of Health Services, A Study of Hospitalization Among North Carolina Children and Youth, Public Health Statistics Branch, July 26, 1974.

(9) See also Samuel D. Kaplan and Albert I. Mendeloff, "PAS Full Coverage Areas: A Resource for Epidemiologic Research," Journal of Chronic Diseases, Vol. 28, 1975, p. 593-599.

Department of Human Resources
Division of Health Services
State Center for Health Statistics
P.O. Box 2091
Raleigh, North Carolina 27602

Bulk Rate
U.S. Postage
PAID
RALEIGH, N.C. 27602
PERMIT NO. 1344